

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
19	<b>87.4</b> (-3.5)	<b>88.5</b> (+1.4)	83.2 (+1.4)	85.3 (+1.6)	77.1 (-1.3)
38	<b>90.2</b> (-3.3)	<b>90.9</b> (+1.0)	<b>87.4</b> (+0.1)	88.5 (-0.6)	81.7 (+0.5)
96	<b>91.7</b> (-0.9)	<b>91.9</b> (-0.5)	<b>90.5</b> (-0.8)	95.0 (-1.5)	85.1 (-0.3)
134	<b>95.2</b> (-2.7)	<b>93.6</b> (-0.8)	<b>93.1</b> (-2.8)	97.7 (-0.9)	91.6 (-2.7)
	<b>Regression</b>				
19	83.7 (+0.5)	87.6 (+3.8)	<b>84.5</b> (+0.9)	91.9 (+7.1)	80.5 (+2.8)
38	86.7 (+0.9)	90.6 (+3.1)	86.8 (+1.0)	91.9 (+3.6)	84.0 (+3.8)
96	91.3 (+0.6)	91.7 (+2.3)	88.1 (+0.2)	94.9 (+1.3)	89.2 (+0.1)
134	93.3 (+0.3)	92.4 (+0.7)	89.9 (+0.6)	<b>98.5</b> (+1.2)	93.4 (+0.3)
	<b>Specialized Regression</b>				
19	83.0 (+1.1)	85.9 (+1.6)	83.7 (+1.6)	<b>92.1</b> (+4.5)	<b>86.5</b> (+4.8)
38	86.8 (+2.1)	86.7 (+1.2)	86.4 (+0.4)	<b>92.3</b> (+1.1)	<b>89.2</b> (+1.8)
96	90.6 (+0.1)	88.9 (+0.3)	87.8 (+0.2)	<b>95.7</b> (+1.7)	<b>93.8</b> (+2.2)
134	93.0 (+0.2)	91.2 (+0.8)	89.9 (+0.7)	98.1 (+1.8)	<b>95.9</b> (+1.5)

Table 1: Decision tree classification accuracy on performance specialization for Apache on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
256	<b>99.9</b> (-0.1)	98.6 (+1.1)	<b>99.9</b> (-0.0)	<b>100.0</b> (-0.0)	<b>100.0</b> (-0.0)
512	<b>99.9</b> (-0.0)	98.9 (+0.1)	<b>99.9</b> (-0.0)	<b>100.0</b> (-0.0)	<b>100.0</b> (-0.0)
1280	<b>100.0</b> (-0.0)	99.3 (+0.3)	<b>99.9</b> (-0.0)	<b>100.0</b> (-0.0)	<b>100.0</b> (-0.0)
1792	<b>100.0</b> (-0.0)	<b>99.4</b> (+0.3)	<b>99.9</b> (-0.0)	<b>100.0</b> (-0.0)	<b>100.0</b> (-0.0)
	<b>Regression</b>				
256	99.8 (+0.0)	<b>98.6</b> (+0.5)	99.9 (+0.0)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
512	99.8 (+0.0)	<b>99.0</b> (+0.6)	99.9 (+0.0)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
1280	99.9 (+0.0)	<b>99.4</b> (+0.5)	99.9 (+0.0)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
1792	100.0 (+0.0)	99.4 (+0.4)	99.9 (+0.0)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
	<b>Specialized Regression</b>				
256	99.8 (+0.0)	97.4 (+0.7)	99.8 (+0.1)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
512	99.8 (+0.1)	98.1 (+0.5)	99.8 (+0.1)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
1280	99.6 (+0.1)	98.6 (+0.4)	99.7 (+0.1)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)
1792	99.7 (+0.2)	98.7 (+0.2)	99.8 (+0.0)	<b>100.0</b> (+0.0)	<b>100.0</b> (+0.0)

Table 2: Decision tree classification accuracy on performance specialization for BerkeleyC on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
18	<b>70.4</b> (+3.9)	<b>89.5</b> (+8.7)	<b>100.0</b> (-0.0)	92.2 (-0.2)	88.6 (-1.6)
36	<b>73.6</b> (-1.6)	<b>91.6</b> (+3.5)	<b>100.0</b> (-0.0)	93.8 (-0.1)	95.9 (-6.6)
90	<b>74.5</b> (-2.2)	<b>92.2</b> (+3.4)	<b>100.0</b> (-0.0)	96.6 (-1.1)	96.5 (-0.1)
125	<b>77.2</b> (-1.7)	<b>93.1</b> (+0.2)	<b>100.0</b> (-0.0)	97.6 (-1.2)	97.1 (+0.0)
	<b>Regression</b>				
18	68.4 (+1.3)	85.8 (+4.0)	<b>100.0</b> (+0.0)	<b>92.5</b> (+1.9)	94.0 (+8.8)
36	69.1 (+2.5)	89.5 (+4.0)	<b>100.0</b> (+0.0)	<b>95.2</b> (+0.6)	96.4 (+2.5)
90	72.1 (+1.2)	92.1 (+0.5)	<b>100.0</b> (+0.0)	96.0 (+0.8)	96.9 (+0.5)
125	75.9 (+1.0)	92.7 (+0.8)	<b>100.0</b> (+0.0)	96.7 (+1.0)	97.4 (+0.6)
	<b>Specialized Regression</b>				
18	70.2 (+3.9)	82.9 (+0.8)	<b>100.0</b> (+0.0)	92.3 (+0.4)	<b>96.4</b> (+6.6)
36	70.3 (+1.7)	83.1 (+1.2)	<b>100.0</b> (+0.0)	94.7 (+1.0)	<b>96.6</b> (+0.2)
90	73.2 (+2.9)	83.0 (+0.3)	<b>100.0</b> (+0.0)	<b>98.2</b> (+0.3)	<b>97.3</b> (+0.5)
125	76.1 (+5.0)	83.4 (+0.1)	<b>100.0</b> (+0.0)	<b>98.7</b> (+0.4)	<b>97.6</b> (+0.6)

Table 3: Decision tree classification accuracy on performance specialization for BerkeleyJ on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
230	<b>83.5</b> (-2.8)	<b>82.5</b> (-0.2)	88.5 (-0.5)	87.4 (+0.6)	87.3 (-1.0)
460	<b>84.4</b> (-0.7)	<b>86.7</b> (-1.2)	90.6 (+0.1)	89.2 (+0.4)	91.4 (-1.1)
1152	<b>88.7</b> (-0.1)	<b>90.3</b> (-0.7)	<b>92.9</b> (-0.2)	92.4 (-0.5)	94.9 (-1.3)
1612	<b>91.3</b> (-1.4)	91.0 (-0.2)	<b>93.9</b> (-0.5)	92.8 (+0.2)	95.0 (-0.9)
	<b>Regression</b>				
230	79.5 (+0.8)	80.6 (+1.5)	88.4 (+0.1)	87.5 (+0.0)	87.4 (+0.3)
460	81.8 (+0.3)	84.3 (+1.9)	90.7 (+0.2)	89.7 (+0.6)	91.6 (+0.3)
1152	86.4 (+0.4)	88.7 (+2.0)	92.7 (+0.1)	91.6 (+0.2)	94.1 (+0.2)
1612	88.0 (+0.0)	89.8 (+2.1)	93.5 (+0.3)	92.4 (+0.4)	94.9 (+0.0)
	<b>Specialized Regression</b>				
230	80.0 (+0.5)	81.6 (+0.7)	<b>88.7</b> (+0.2)	<b>90.1</b> (+0.3)	<b>90.8</b> (+0.1)
460	84.3 (+1.5)	85.0 (+0.4)	<b>90.7</b> (+0.2)	<b>91.5</b> (+0.3)	<b>93.5</b> (+0.4)
1152	87.7 (+0.9)	89.5 (+0.3)	92.8 (+0.1)	<b>94.0</b> (+0.3)	<b>96.2</b> (+0.3)
1612	88.8 (+0.2)	<b>91.1</b> (+0.6)	93.2 (+0.0)	<b>94.8</b> (+0.3)	<b>96.8</b> (+0.1)

Table 4: Decision tree classification accuracy on performance specialization for Dune on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
1348	<b>91.2</b> (+1.4)	93.2 (+1.6)	<b>96.3</b> (+1.2)	93.9 (+1.7)	95.6 (+0.6)
2697	<b>94.4</b> (+3.2)	<b>95.2</b> (+1.3)	<b>97.7</b> (+0.6)	95.3 (+0.6)	96.6 (+0.2)
6742	<b>96.6</b> (+2.3)	<b>96.4</b> (+1.1)	<b>98.1</b> (-0.1)	96.3 (+0.1)	97.0 (-0.1)
9439	<b>97.3</b> (+1.9)	<b>96.4</b> (+0.6)	<b>98.2</b> (-0.1)	96.8 (-0.2)	97.1 (+0.0)
	<b>Regression</b>				
1348	87.0 (+1.2)	<b>93.3</b> (+1.0)	96.1 (+1.3)	94.2 (+1.5)	96.1 (+0.8)
2697	88.9 (+1.3)	94.3 (+0.1)	97.6 (+0.5)	95.3 (+0.6)	96.9 (+0.3)
6742	91.2 (+0.7)	95.3 (+0.1)	98.0 (+0.0)	95.2 (+0.0)	97.2 (+0.0)
9439	92.2 (+0.2)	95.2 (+0.1)	98.1 (+0.1)	95.2 (+0.3)	97.2 (+0.1)
	<b>Specialized Regression</b>				
1348	87.0 (+0.3)	89.2 (+1.5)	96.0 (+1.1)	<b>95.0</b> (+1.2)	<b>97.1</b> (+0.4)
2697	88.9 (+0.2)	90.1 (+0.7)	97.1 (+0.2)	<b>96.3</b> (+0.4)	<b>98.0</b> (+0.3)
6742	91.2 (+0.1)	88.9 (+0.3)	97.8 (+0.0)	<b>97.0</b> (+0.1)	<b>98.6</b> (+0.3)
9439	91.8 (+0.1)	89.1 (+0.3)	97.8 (+0.0)	<b>97.2</b> (+0.0)	<b>98.8</b> (+0.2)

Table 5: Decision tree classification accuracy on performance specialization for HIPAcc on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
345	<b>91.8</b> (+4.3)	<b>91.1</b> (+2.6)	91.3 (+0.8)	96.7 (+0.4)	96.7 (+0.3)
691	<b>92.8</b> (+2.9)	92.0 (+1.6)	92.7 (-0.1)	97.7 (+0.5)	98.3 (+0.3)
1728	93.7 (+1.1)	94.5 (+1.3)	94.7 (+0.6)	<b>98.7</b> (-0.1)	99.4 (+0.4)
2419	<b>94.8</b> (+0.6)	95.2 (+0.6)	95.4 (+0.4)	<b>99.2</b> (+0.4)	99.7 (+0.2)
	<b>Regression</b>				
345	90.1 (+2.4)	91.1 (+1.7)	<b>91.8</b> (+0.8)	97.0 (+0.9)	97.6 (+0.9)
691	91.6 (+1.6)	<b>92.7</b> (+1.3)	<b>93.3</b> (+0.2)	97.8 (+0.6)	<b>98.7</b> (+0.6)
1728	<b>94.1</b> (+0.6)	<b>95.2</b> (+0.7)	<b>95.5</b> (+0.3)	98.6 (+0.4)	99.4 (+0.3)
2419	94.8 (+0.6)	<b>96.0</b> (+0.7)	<b>96.3</b> (+0.3)	99.0 (+0.5)	99.6 (+0.1)
	<b>Specialized Regression</b>				
345	88.8 (+1.7)	89.8 (+1.4)	91.3 (+0.8)	<b>97.4</b> (+0.7)	<b>97.9</b> (+0.6)
691	90.8 (+1.5)	92.0 (+1.4)	93.1 (+0.6)	<b>97.9</b> (+0.7)	98.7 (+0.5)
1728	93.3 (+1.0)	94.6 (+1.0)	95.0 (+0.6)	98.7 (+0.3)	<b>99.6</b> (+0.3)
2419	94.7 (+1.1)	95.6 (+0.7)	95.8 (+0.4)	99.1 (+0.4)	<b>99.8</b> (+0.1)

Table 6: Decision tree classification accuracy on performance specialization for HMSGP on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
102	<b>95.2</b> (-0.2)	<b>95.9</b> (-0.1)	<b>94.4</b> (+0.8)	80.3 (+5.4)	92.3 (+8.0)
204	<b>94.7</b> (-1.9)	<b>96.2</b> (-0.1)	<b>94.8</b> (+0.2)	80.5 (+1.7)	94.9 (+10.3)
512	<b>96.0</b> (-0.3)	<b>97.0</b> (-0.5)	<b>95.5</b> (-0.0)	82.7 (+0.5)	95.6 (+7.9)
716	<b>96.3</b> (-0.5)	<b>97.2</b> (-0.1)	<b>96.3</b> (-0.0)	84.9 (+0.3)	95.9 (+1.3)
	<b>Regression</b>				
102	94.8 (+0.6)	95.8 (+0.2)	94.2 (+1.2)	80.4 (+3.8)	88.1 (+7.8)
204	93.6 (+0.1)	95.8 (+0.3)	94.6 (+1.1)	80.6 (+2.4)	93.1 (+10.2)
512	95.4 (+0.2)	96.1 (+0.2)	95.2 (+0.5)	83.7 (+0.6)	92.1 (+5.5)
716	95.9 (+0.3)	96.4 (+0.2)	95.7 (+0.4)	85.2 (+1.4)	93.2 (+4.8)
	<b>Specialized Regression</b>				
102	90.6 (+0.9)	95.0 (+0.5)	93.1 (+1.0)	<b>84.0</b> (+4.0)	<b>94.8</b> (+3.6)
204	90.9 (+0.1)	95.0 (+0.2)	93.7 (+0.5)	<b>85.0</b> (+1.8)	<b>95.3</b> (+3.4)
512	91.9 (+0.4)	96.0 (+0.5)	94.7 (+0.0)	<b>85.9</b> (+0.7)	<b>95.9</b> (+1.1)
716	93.1 (+0.3)	96.4 (+0.6)	95.4 (+0.2)	<b>86.9</b> (+0.1)	<b>96.6</b> (+1.0)

Table 7: Decision tree classification accuracy on performance specialization for LLVM on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
455	<b>59.2</b> (-0.7)	<b>63.5</b> (+1.6)	70.8 (+2.3)	72.4 (+1.2)	80.9 (+1.3)
910	<b>59.5</b> (-0.4)	<b>65.1</b> (+0.4)	72.7 (+2.0)	73.4 (+0.6)	82.1 (+0.5)
2276	<b>61.0</b> (+0.2)	<b>67.1</b> (-0.5)	75.5 (+0.8)	76.5 (-0.0)	84.3 (-0.9)
3187	<b>61.6</b> (-0.0)	<b>67.4</b> (-0.6)	76.5 (+0.1)	78.8 (-1.1)	84.2 (+0.1)
	<b>Regression</b>				
455	57.7 (+0.3)	62.2 (+2.0)	<b>71.1</b> (+2.8)	72.6 (+1.6)	74.7 (+3.2)
910	58.7 (+0.4)	62.5 (+0.5)	<b>73.6</b> (+2.4)	73.6 (+0.2)	75.2 (+2.3)
2276	58.4 (+0.2)	64.1 (+0.1)	<b>76.0</b> (+1.0)	75.3 (+0.3)	76.2 (+1.0)
3187	59.1 (+0.2)	64.9 (+0.8)	<b>76.6</b> (+0.2)	76.0 (+0.6)	76.9 (+0.7)
	<b>Specialized Regression</b>				
455	57.4 (+0.8)	60.0 (+0.1)	65.6 (+0.9)	<b>76.1</b> (+2.3)	<b>82.6</b> (+2.5)
910	58.1 (+0.0)	61.1 (+0.3)	66.6 (+0.4)	<b>77.7</b> (+0.7)	<b>84.2</b> (+1.0)
2276	57.6 (+0.6)	60.3 (+0.6)	67.2 (+0.4)	<b>80.4</b> (+0.2)	<b>86.4</b> (+0.1)
3187	57.1 (+0.8)	59.4 (+0.5)	67.7 (+0.9)	<b>81.2</b> (+0.3)	<b>87.6</b> (+0.3)

Table 8: Decision tree classification accuracy on performance specialization for SQLite on three strategies. Bold represents the best result among other strategies including feature selection, the value in brackets is the difference made by feature selection

Training set size	Acceptable configurations				
	10%	20%	50%	80%	90%
	<b>Classification</b>				
9256	84.4% (+1.7)	88.4% (+0.2)	90.3% (+0.2)	91.8% (+0.7)	91.6% (+2.8)
18512	85.4% (+0.4)	89.0% (+0.4)	91.1% (-0.0)	92.5% (+1.0)	92.4% (+0.2)
46281	<b>87.3%</b> (-0.4)	<b>90.1%</b> (-0.2)	92.1% (+0.2)	93.1% (+0.4)	93.5% (+1.2)
64793	<b>87.4%</b> (-0.4)	<b>89.9%</b> (+0.0)	92.6% (-0.2)	93.4% (+0.4)	93.7% (+1.3)
	<b>Regression</b>				
9256	<b>85.1%</b> (+1.8)	<b>88.5%</b> (+1.4)	<b>91.6%</b> (+1.3)	92.0% (+0.8)	92.0% (+2.4)
18512	<b>86.1%</b> (+1.9)	<b>89.5%</b> (+1.2)	<b>92.0%</b> (+0.3)	92.9% (+1.0)	92.4% (+0.8)
46281	86.8% (-1.0)	89.7% (+0.0)	<b>92.9%</b> (+0.7)	93.8% (+0.6)	94.2% (+1.2)
64793	86.8% (-0.5)	89.9% (+0.2)	<b>92.9%</b> (-0.1)	93.9% (+0.2)	94.1% (+1.0)
	<b>Specialized Regression</b>				
9256	85.0% (+2.0)	87.2% (+1.1)	91.5% (+1.0)	<b>94.8%</b> (+0.7)	<b>95.7%</b> (+0.5)
18512	84.4% (+0.6)	87.6% (-0.0)	91.5% (+0.8)	<b>95.1%</b> (+0.7)	<b>96.3%</b> (+0.6)
46281	84.7% (+0.9)	87.8% (-0.1)	92.3% (+0.4)	<b>95.6%</b> (+0.7)	<b>96.8%</b> (+0.7)
64793	86.2% (-0.3)	89.1% (-0.4)	92.8% (+0.5)	<b>95.8%</b> (+0.6)	<b>97.0%</b> (+0.6)

Table 9: Decision tree classification accuracy on performance specialization for Linux kernel on three strategies. Bold represents the best result among other strategies (including feature selection), the value in brackets is the difference made by feature selection